R E P O R T R E S U M E S

ED 002 973 24

PROBLEM-SOLVING PROFICIENCY AMONG ELEMENTARY SCHOOL TEACHERS, I. THE DEVELOPMENT OF CRITERIA.

BY- TURNER, RICHARD L. FATTU, NICHOLAS A.
INDIANA UNIV., BLOOMINGTON, INST. OF EDUC. RES.

REPORT NUMBER CRP-419-1 PUB DATE MAY 60
CONTRACT OEC-SAE-7790
EDRS PRICE MF-$0.18 HC-$2.76 69P.

DESCRIPTORS- *PROBLEM SOLVING, *TEACHING QUALITY, ELEMENTARY SCHOOL TEACHERS, *TEACHER EVALUATION, EFFECTIVE TEACHING, *CONCEPT FORMATION, TEACHER QUALIFICATIONS, *BEHAVIOR PATTERNS, BEHAVIOR STANDARDS, EDUCATIONAL OBJECTIVES, BLOOMINGTON, INDIANA

THE OBJECTIVE OF THIS STUDY WAS TO GATHER AND ANALYZE DATA ABOUT THE PROBLEM-SOLVING PROCESSES OF ELEMENTARY SCHOOL TEACHERS IN ORDER TO DEVELOP AN OPERATIONALLY DEFINED CONCEPT ON THE CRITERIA FOR PROFESSIONAL, HIGH PROFICIENCY QUALITIES IN TEACHERS. THREE QUESTIONS WERE ASKED--(1) WHAT IS AN ELEMENTARY SCHOOL TEACHER, (2) WHAT IS PROBLEM-SOLVING PROFICIENCY, AND (3) IS PROBLEM-SOLVING PROFICIENCY A USEFUL CONCEPT IN STUDYING TEACHER BEHAVIOR. TEACHERS WERE DEFINED AS PROFESSIONAL PERSONS WHO MUST ASSUME AND BE CAPABLE OF ATTAINING CERTAIN EDUCATIONAL OBJECTIVES DESIRED BY LEADERS IN THE PROFESSION. TYPES OF CRITERIA WERE THEN SET UP ACCORDING TO PROBLEM-SOLVING PROFICIENCY. THESE WERE THEN USED IN OUTLINING A PROBLEM-SOLVING PROFICIENCY CONCEPT AND FOR DEVELOPING A DESCRIPTION OF THAT CONCEPT TO DETERMINE THE USEFULNESS OF THE CONCEPT. IT WAS NECESSARY TO IDENTIFY A GROUP OF ELEMENTARY TEACHERS WHO ENJOYED A HIGH LEVEL OF SKILL IN DEFINING AND RESOLVING TEACHING PROBLEMS AND WHO SHOWED IMPROVEMENT IN DEVELOPING THIS SKILL DURING THEIR TENURE AS CLASSROOM AND LABORATORY TEACHERS. IT THEN BECAME POSSIBLE TO CONDUCT EMPIRICAL STUDIES FOR DETERMINING WHETHER OR NOT THE CONCEPT OF PROBLEM-SOLVING PROFICIENCY WAS A SIGNIFICANT CONCEPT. THE CURRENT STUDY STOPS HERE. DEVELOPED CRITERIA WERE APPLIED TO TEACHER SAMPLES IN THE AREAS OF ARITHMETIC AND READING INSTRUCTION. THESE PROCEDURES AND THEIR OUTCOMES ARE DISCUSSED IN ADDITIONAL REPORTS PREPARED UNDER COOPERATIVE RESEARCH PROJECT 419. (JH)

ED002973

REFERENCE COPY

COOPERATIVE RESEARCH PROJECT

No. 429-I

419

MONOGRAPH

of the

INSTITUTE of EDUCATIONAL RESEARCH

at

INDIANA UNIVERSITY

Problem Solving Proficiency Among Elementary School Teachers

I. The Development of Criteria

by

RICHARD L. TURNER
NICHOLAS A. FATTU

MAY 1960

PROBLEM SOLVING PROFICIENCY AMONG ELEMENTARY SCHOOL TEACHERS

I. THE DEVELOPMENT OF CRITERIA

by

Richard L. Turner

Nicholas A. Fattu

The research on which this study was based was supported in part by the Cooperative Research Program of the U. S. Office of Education under Project 419(7790).

Published by the
INSTITUTE OF EDUCATIONAL RESEARCH
School of Education
Indiana University

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Price, $1.00, postpaid
For sale by the Institute of Educational Research
School of Education
Bloomington, Indiana

ACKNOWLEDGEMENTS

For permission to quote the authors make grateful acknowledgement to the following sources: John Wiley and Sons, Inc. for a quotation from Bruner, Goodnow and Austin's, A Study of Thinking (Copyright 1956); Prentice-Hall, Inc. for quotations from Lieberman, Education as a Profession (Copyright, 1956); the American Educational Research Association for a quotation from the Review of Educational Research (June, 1952).

The authors gratefully acknowledge the support of the Cooperative Research Program of the United States Office of Education, Department of Health, Education, and Welfare, for the project of which this report is a part.

TABLE OF CONTENTS

	Page
INTRODUCTION	1
Origins of the Rationale	1
PROBLEM 1. What is an Elementary School Teacher?	4
The Teacher as a Professional Person	4
Defining the Function of Elementary School Teachers on the Basis of Instructional Objectives	5
Defining Function on the Basis of Problem Solving.	6
Defining a Problem in Teaching	7
Controlling Teaching Problems	12
Summary, Problem 1	13
PROBLEM 2. What is Problem Solving Proficiency?	14
Procedure 1. Recourse to an Ultimate Criterion	15
Procedure 2. Test by Consistency	21
Procedure 3. Test by Consensus	25
Categorizing Procedures and Operational Definitions of Problem Solution	26
Operationally Defining Proficiency	29
Summary, Problem 2	34

Page

PROBLEM 3. Is Problem Solving Proficiency a Useful Concept in Studying Teacher Behavior? 36

The Validation of Laboratory Problems 38

Experimental Designs for Validation on a Transfer Criterion 40

Establishing Additional Criteria for Intrinsic Validity 43

Establishing the Concept of a Professional Teacher 46

SUMMARY 53

LIST OF FIGURES

	Page
FIGURE 1. An Illustrative Problem in Teaching Arithmetic .	10
FIGURE 2. A Summary of Criteria for Problem Solution . .	35a

INTRODUCTION

The present monograph is the first report in a series of studies dealing with problem solving proficiency as a central concept through which to gain better understanding of skill in teaching in the elementary school. The function of this monograph is to set forth in concise outline the systematic rationale by which the series of studies are guided, rather than to report procedures and data on problem solving proficiency in various areas of teaching. The latter is the function of subsequent reports. A more comprehensive discussion of the background to the problem, and particularly the relationship between the concept of "teacher effectiveness" and the concepts set forth in the present monograph may be found in The Bulletin of the Indiana University School of Education, May, 1960. The present discussion may be best understood if this Bulletin is read first.

Origins of the rationale.

The interest of the investigators in problem solving proficiency relative to elementary school teaching arose from the proposition that problem solving proficiency is a way of speaking of the quality of the transactions an individual has with his environment, and particularly a way of speaking about the competence an individual displays in dealing with an environment which involves many "novel" (or variant) situations.

It is impossible to specify what a "good elementary school teacher" is by purely empirical procedures, since value judgments are implicitly involved, but it is none the less clear that the intensional or extended meaning of the term "good elementary school teacher" involves both the notion that the teacher can define and resolve the problems which confront him upon a given occasion and that he improves in problem solving skill during his tenure in the classroom. These notions will be discussed in various perspectives throughout the paper and ultimately the concept of a "professional teacher", differing from the concept of a "good teacher" in several respects, will be set forth. However, in order to give these notions content, it is first necessary to set forth a working purpose for the projected series of studies which have grown out of the concept of problem solving proficiency. This purpose will later be differentiated and more fully explicated, but will serve to get matters under way at this point.

The purpose of the projected studies exploring problem solving proficiency relative to elementary school teachers is to identify those teachers who (i) enjoy a high level of skill in defining and resolving those problems with which elementary school teachers deal as an implicit part of their function and (ii) have improved in this skill during their tenure as classroom teachers.

Implicit when this working purpose is assumed, are three clusters of problems: 1) to define what an elementary school teacher is, 2) to define what problem solving proficiency is, and 3) to determine whether problem solving proficiency is a useful concept in the study of teacher behavior and teaching skill.

PROBLEM I

What Is An Elementary School Teacher

As Barr (1, p. 261) has pointed out, both teachers and teaching can be examined from many points of view. The teacher may be viewed as a transmitter of knowledge, guider or director of learning activities, substitute parent, counselor, adult model, "producer of effects", staff member, community member, and member of a profession. Teaching may similarly be viewed in many perspectives. Which perspective of teachers and teaching one chooses to take will determine the kind of evidence collected and the type of assertions about, or interpretations of, teaching and teacher behavior.

Out of all the possible ways of viewing teachers and teaching, the authors have chosen two. One perspective deals with the teacher as a member of a profession; the other with the teacher as a solver of diverse classes of problems. These two perspectives are conjunctive.

The teacher as a professional person.

From Lieberman's discussion of the characteristics of a profession, professional personnel may be described as (i) performing an essential social service, (ii) placing an emphasis on intellectual rather than physical techniques, particularly on defining and collecting evidence to resolve problems, (iii) having had a long period of specialized training, (iv) having a broad range of autonomy, (v) accepting broad responsibility

for personal judgments, (vi) placing an emphasis on service rather than economic gain, and (vii) belonging to a comprehensive self-governing body of practitioners (13, pp. 2-6).* Of all these descriptive characteristics of professional personnel, the authors focus upon the first and second.

Defining the function of elementary school teachers on the basis of instructional objectives.

The social service performed by the members of a profession may be defined in several parts. It is defined by what the practitioners do, by social expectation, and by the goals or objectives the professional leaders and practitioners set. As Lieberman suggests, when the function of a profession is not clearly defined either by the public or by the members of the profession, issues bearing on professional education, professional ethics, remuneration and the like, arise and the status of the profession is held in doubt (13, p. 2). Elementary school teaching appears to be in this position with regard to the definition of its functions, insofar as there is considerable variation in what local populations expect from such teachers. However, the past several years have yielded publications (3, 8, 10, 12) which indicate that within the profession there is increasing interest in and agreement on the objectives of teaching in the elementary school.

*Reprinted with permission from Myron Lieberman, Education as a Profession, (C) 1956 by Prentice-Hall, Inc., Englewood Cliffs, N.J.

The problem of defining the function of elementary school teachers may be in part resolved by assuming that such statements of objectives as those referred to explicate the function to be discharged sufficiently well to be used (i) for purposes of defining the domain within which elementary school teachers work, and (ii) as working criteria for what elementary school teachers ought to be able to do. When these objectives are used as working criteria, it is in effect being asserted that it is desirable that elementary school teachers be capable of bringing these objectives about in the behavioral repertoire of pupils.

Defining function on the basis of problem solving.

The objectives sought in the elementary school, as outlined by such persons as Kearney, Brownell, Havighurst, and Greene and Gray define, although not exhaustively, the domain in which it is desirable that elementary school teachers work. Referring again to Lieberman's characteristics of a profession, elementary school teachers might be expected to use intellectual techniques to define problems and collect information with which to resolve them within the domain defined by the objectives previously cited. To the question "what do elementary school teachers do?", then, the answer would be "define and attempt to resolve problems within the domain delimited by their objectives." One purpose of this answer is heuristic and is directed toward establishing a way of looking at teachers and teaching which diverges some-

what from the ways in which these matters have been viewed before. A second purpose is to establish relationships between teaching behavior and problem solving behavior.

Defining a problem in teaching.

If teaching consists of various phases of problem solving, it may be asked how the investigators conceptualize problems in teaching. The general answer to this question is that a problem arises in teaching each time a teacher assumes, with or without verbalization, that a particular goal or objective--a particular behavioral outcome--is to be brought into concretion in one or more pupils through the instrumental responding of the teacher. Or, to state the answer in slightly different terms, a problem arises in teaching whenever there is a discrepancy between the actual behavior of a pupil and the behavior which a teacher holds to be desirable for that pupil.

Beyond this general definition of a problem in teaching, the authors conceptualize teaching problems to be analyzable into parts or segments. These segments are called tasks. Some types of tasks exist by strict implication from the assumption that teaching problems arise because teachers assume that certain outcomes can be brought about in pupils. One such type of task consists of goal setting by the teacher. Goal setting may or may not occur directly in the classroom. It may, for instance, occur at home when the teacher is preparing for his classes.

It is not observable unless the teacher verbalizes the goals he is setting and even then subsequent behavior may cast doubt on whether the teacher's verbal behavior coincides with his other observable behavior.

The goal setting behavior of teachers might be studied strictly from the viewpoint of the individual teacher, i.e. by examining only those goals a teacher sets with no reference to other goals. It is apparent, however, that the task of goal setting to an individual teacher is the task of explicating what his function as a teacher is, and therefore is an instance of the class "teacher functions" and is related in some degree to the function of a teacher as explicated by other persons, be they leaders in the profession or the administrative officers in the school system in which the teacher teaches. The rather significant implications of this notion are taken up in the final section of this paper.

A second type of task is closely related to the first and deals with the diagnosis of pupil difficulties. Diagnostic tasks obviously involve evaluation, i.e. the assessment of pupil status, but go beyond evaluation in that the particular phase or aspect of the goal which the pupil has not achieved must be isolated, his errors en route to the goal must be examined and then a choice among remedial procedures, or additional instrumental responses by the teacher, must be made. From the viewpoint of the authors, the diagnostic task is the epitome of teaching.

First, a diagnostic task involves problem definition, i.e. isolation of exactly that part of the goal assumed by the teacher which is not fully achieved by the pupil. The goal might be a subject-matter goal such as adding by endings in column addition, or a social goal such as taking turns in playground games. Second, after the problem is defined, the teacher is confronted with what in maze learning could be called a choice point, or in problem solving, a set of alternatives. The problem is, of course, to choose the alternative which will most nearly satisfy the conditions of the goal as determined by the definition of the problem. Whatever alternative the teacher chooses will involve the performance of other tasks the nature of which is contingent on the type of goal the teacher is attempting to bring about.

In solving problems in content areas or in the common cognitive skills, the tasks of the teacher which follow goal setting or diagnosis will frequently include a task which centrally involves the selection of materials by which the responses leading to or constituting the goal state for the pupil can be elicited. These materials might range from concrete materials in arithmetic to field trips and films in science and social studies through various kinds of written material in the common cognitive skills, up to questions or explanations from the teacher. The utilization of such materials is usually preceded by a task which involves providing a set for pupils, so that they are ready to respond to the material, and finally a task involving the administration of reinforcement by the teacher or else providing for reinforcement for the pupil from

some other aspect of the situation. A problem of a comparatively simple nature is shown on Figure 1.

	What the teacher does	What the teacher does (conceptualized as a task)
PROBLEM SOLVING	John should respond "9" to the stimulus 3x3, "6" to 3x2, "12" to 3x4.	First task--goal setting
	Cuts paper markers for John to manipulate to group and regroup quantities above.	Second task--selecting materials
	Orients John toward the goal.	Third task--establishing a set
	Presents markers to John, explains grouping, presents written arithmetic facts, physically helps John to attain correct groupings.	Fourth task--utilization of materials
	Verbal presentation of arithmetic facts with encouragement of oral responding, correction or praise of responses.	Fifth task--eliciting and reinforcing responses
	Test of responses--1. on criterion of automatic recall, 2. on criterion of use in problem solution.	Sixth task--assessing behavioral outcomes.

Figure 1. An Illustrative Problem in Teaching Arithmetic.

It is apparent from an examination of Figure 1 that the notion of

a task is not a hard and fast one. Where the fourth task leaves off and the fifth task or even the sixth task begins is difficult to specify. Nevertheless, in examining the tasks in sequence it may be observed that each change of task or each segment of the problem involves a change in the locus of the teacher's activity, including his verbal activity, in comparison to the preceding tasks. For instance, in the second task the locus of the teacher's activity may be spatially centered at the supply cabinet, at the paper cutter, and at his desk, while temporally it may occur after school or during a free period. In the third task the spatial locus of activity shifts, and centers on the pupil while the locus of his verbal behavior centers around words that arouse and direct the pupil. In the fourth task the locus of teacher activity changes from arousing and directing or using arousing and directing language, to using explanatory language and to the manipulation of materials. In task five the locus of verbal behavior again shifts, this time towards approving behavior, including such behavior as "good", "fine", "you're doing well", "that's right", and so on. It may thus be anticipated that each time the locus of a teacher's activity changes he is considering or dealing with some phase or segment of an ongoing problem. In the classroom, of course, a teacher may have many problems in process at a given time so that it is difficult to sort out which teacher activity is related to which ongoing problem.

Now, while comparatively simple problems such as that suggested

in Figure 1 can become rather complex, the complexity of the problems a teacher faces can be rapidly compounded by two basic means. One means is by increasing the number of pupils with which the teacher has to deal. At some point goal setting can no longer be done on an individual basis, but rather must be done on a group basis. When this is the case, the task of goal setting becomes intertwined with the task of grouping pupils and controlling such groups. A second means of increasing the complexity of a problem is to increase the number of types of goals which the teacher sets for the pupil. As the number of goals a teacher sets for a pupil increases, the number of alternative ways he has of proceeding to all goals simultaneously appears to be correspondingly reduced. For instance, independence in reading is a common, general goal which consists of a number of sub-goals or objectives of which one is independence in word attack, characteristically involving the learning of skills in structural and phonetic analysis, while the other is independently going to the bookshelf or library and selecting books at an appropriate level of difficulty for reading for information or pleasure. The problem is to teach word attack skills by procedures which do not condition pupils to avoid reading books.

Controlling teaching problems.

By assuming that teaching involves much problem solving behavior and that teaching problems are composed of segments or tasks, the possibility arose that the problem solving behavior of the teacher might be

studied outside the classroom if problems and tasks similar to those encountered in the classroom were constructed. The advantages in the control of variables in constructed problems performed under laboratory conditions are immediately apparent. The authors thus chose to use test and laboratory situations in the projected studies, rather than study problem solving behavior in the classroom. This choice, however, generated three problems. The first was to develop test and laboratory problems. In view of the complexity of elementary school teaching, problem development is a continuous phase of the studies rather than an initial and terminating one. The second problem was to establish criteria for deciding when a problem is solved. The third problem is to show that constructed test and laboratory problems are valid, i.e. useful in talking about teachers and teaching. How these latter two problems may be solved is discussed in the remaining sections of this paper.

Summary, Problem 1

Teaching and teachers may be examined from many points of view. The present study conceptualizes elementary school teachers as professional persons whose functions may be defined by currently published objectives for the elementary school and by the "problem solving activity" or task performance in which teachers engage in discharging these functions. This conceptualization in no way explains teaching, but rather

provides a perspective within which to examine teaching. Whether utilizing this perspective will be fruitful for establishing systematic knowledge about teacher behavior and for predicting teacher behavior under specific conditions is an unanswered empirical question.

PROBLEM 2

What Is Problem Solving Proficiency?

Initially the question may be divided into two parts (1) what does it mean to solve a problem? (2) what does it mean to be "proficient" as a problem solver?

The first question may be answered with relative ease as long as the solution to the problem is unmistakable and there is at least one known sequence of operations which unequivocably will lead to this solution.* For instance in the PSI apparatus (11), a fixed sequence of operations in a given problem will activate a light which in turn signals problem solution. In this case there is no doubt about problem solution. However it is precisely the absence of a fixed sequence of operations and an unmistakable solution which appears to characterize many problems in elementary school teaching. It is not only a matter of having several possible sets of instrumental responses, but also of having

---

*This is the case only when the problem exists within a closed logical system.

uncertain criteria for problem solution. For instance, if the objective is to have pupils "understand" long division, does one take as the criterion for having attained this objective the ability to solve division problems accurately by the standard algorism, the ability to utilize division to solve practical problems, the ability to verbalize the principles of regrouping in division and the meanings of the components of the problem, the ability to solve division problems using the binary system, or are still other criteria to be applied? Suppose that Mary shows verbal understanding of division this year, but cannot work division examples accurately next year, what then?

In order to deal with problems or tasks in which there are varying degrees of indeterminacy in both the sequence of instrumental responses and problem solution, it is necessary to consider several types of criteria in working toward an operational definition of "problem solving" and "proficiency." The types of criteria considered by the authors are set forth by Bruner, Goodnow and Austin in A Study of Thinking (4, pp. 17-20), and were initially used to describe the procedures or "tests" by which people appear to validate the categorization of perceptual events.

Procedure 1. Recourse to an ultimate criterion.

Bruner, Goodnow and Austin (4, p. 17) illustrate this procedure by use of the mushroom fancier who, out in the woods, might make the ultimate test of whether or not a particular class of mushroom is edible

by nibbling one. This procedure will either lead to illness or will not do so. If it does not do so, then the categorization of the mushroom as "edible" will be correct. In general, whenever the procedure used leads to an unmistakable consequence, this consequence can be used as the ultimate criterion.

The use of ultimate criteria has indeed been attempted in assessing whether teachers resolve their problems. Studies of teacher effectiveness have characteristically viewed effects on pupils as the ultimate test of "effectiveness." As stated in the first Report of the Committee on the Criteria of Teacher Effectiveness:*

> "Effects on pupils have long been accepted as relevant criterion dimensions of teacher effectiveness; pupil growth, change, gain, development, learning and the like, have often been denoted as the ultimate criteria of teacher effectiveness." (14, p. 244).

Effects on pupils is then further defined as follows:

> "The effects on pupils that are relevant as criterion dimensions of teacher effectiveness are the extent to which educational objectives are attained." (14, p. 245)

Since to "attain educational objectives" is a synonym for the "the goal state" or "problem solution," it is apparent that to solve teaching problems and to "produce effects" are the same thing when effects are

---

*Reprinted with permission from the Editor of the Review of Educational Research.

defined as the ultimate criteria.

It is tautological to assert that effects on pupils is an ultimate criterion of teacher effectiveness, but the crucial matter is not simply to accept the tautology; it is first necessary to be able to carry out those procedures which (i) show that a given set of instrumental responses will account for a given effect and (ii) operationally define what the "effect" is, and second to show that carrying out these procedures will lead to useful knowledge about teaching. Considerable doubt may be entertained on some of these points.

To demonstrate that a given sequence of instrumental responses accounts for a given effect is a perennial problem in teacher effectiveness research (1, p. 256; 14, p. 257 ff.). The resolution to that part of the problem which requires adequate control of all variables except the instrumental responses of the teacher, seems to be largely problematic and involves no logical difficulties. However, resolution of that part of the problem which deals with operationally defining effects appears to be a somewhat different matter. The problem in this respect lies in the way "effects" are conceptualized.

The first Report of the Committee on the Criteria of Teacher Effectiveness recognized the problem by asserting that the criterion of effects may be conceptualized on an "ultimacy-proximacy" basis. According to this committee the more "ultimate" criteria, in decreasing order of "ultimacy" are as follows:

"Pupils' achievement and success in life
pupils' achievement in subsequent schooling
pupils' achievement of current educational objectives
pupils' satisfaction with the teacher
parents' satisfaction with the teacher
superintendents' satisfaction with the teacher
..." (14, p. 243).

In examining the first two or "more ultimate" criteria it is immediately recognized that the effects in question cannot be operationally defined at the end of instrumental responding or problem solving because they are implicitly contingent upon future states of affairs. It is no doubt true that the learnings a pupil acquired as a result of the operations of his fourth grade teacher are related to his achievement in grade six, but how they are related depends upon what happened to him in grade five and what is happening to him in grade six. They depend upon what happens in subsequent grades because a transfer of training criterion for the assessment of effects has been assumed, and transfer is implicitly defined by the occurrence of some state of affairs subsequent to original training. The problem is, then, whether a problem in teaching should be considered "solved" or an "effect" considered to have occurred when the training trials are completed, i.e. at the end of the sequence of instrumental responses by the teacher, or whether the problem may be considered solved only when there is some evidence that what was learned as a result of original training transfers.

At this point, one is confronted with a paradox. The procedures which define the ultimate or transfer criterion of teacher effectiveness cannot be carried out until (i) a confusing array of intervening events has occurred and (ii) so much time has passed that there is no longer much point in finding out whether the teacher is effective. On the other hand, to assess effects at the termination of instrumental responding by the teacher eliminates from consideration (i) the possibility that some teachers may produce few but enduring changes in behavior, and (ii) the possibility that some teachers may produce few changes in behavior but produce them in such a way that the acquisition of future desirable behavior is greatly facilitated. As long as the latter possibilities exist, the measurement of effects at the end of the teacher's instrumental responding may be regarded as a convenient and operable criterion for whether a problem is solved, but it cannot be regarded as an ultimate criterion. The fact that an operationally definable criterion of effects is not an ultimate criterion leaves its status as a criterion of problem solution indeterminate. To say that it is a possible criterion is indeed true, but perhaps the more focal question is whether it is a useful criterion.

The second question is whether utilization of the criterion of effects will lead to useful knowledge. In the context of the present study it may be asked whether the criterion of effects is useful for studying the level

of problem solving proficiency and improvements in such proficiency among elementary teachers. The answer to this question is that the criterion of effects is of very limited utility. There are a number of reasons for this answer. First, if problem solving is studied in the classroom there is no way to control the problems which the teacher attacks. This means that there is no basis for comparing either the instrumental responses of or the effects produced by one teacher to those of another teacher. It is the case, then, that research data would be composed of the protocols of many teachers, each of whom had responded to different stimuli. Such data would be useful for discussing the problems a particular teacher encountered and how he responded, but the data would have no generalizability and would be useless for making predictions about teachers as a class unless the situations to which the teachers responded were functionally identical. Second, if there is no means by which to control problems, it is ipso facto impossible to show that a teacher has either increased or decreased in proficiency since there is no common basis on which to compare a teacher's performance at one time to her performance at another. Third, it must be possible not only to examine the effects of instrumental responding, but also the instrumental responding itself, if understanding of the problem solving behavior of the teacher is to follow. Just how the instrumental responses of a teacher can be isolated in the classroom is a procedural problem which, up to this point, has

not been resolved, probably because the goals toward which the instrumental responses of the teacher are directed are difficult to infer from the observation of his behavior. However, if this procedural problem could be resolved, it would be possible to study the relationship between problem solving activity in the classroom, including the extent to which such activity leads to given effects, and such activity under laboratory conditions.

Procedure 2. Test by consistency.

What Bruner and his collaborators mean by test or validation by consistency is set forth in the following quotation:*

> Validation by consistency is perhaps nowhere better illustrated than in modern taxonomic research. At the lowest level, so-called alpha taxonomy, one seeks to differentiate as many species as possible in terms of whatever defining properties are visible. Such a technique leads to vast multiplication of categories. There are now about a third of a million species of plants, and each year about 5,000 are added. There are about 2 million species and sub-species of animals and it is estimated that new ones are being added at the rate of 10,000 per year. Some estimates of the number of insect species run as high as 3 million (Silverstri, 1929), and given new radiological methods of producing mutations, the possible number of virus types seems almost unlimited. As far as identification is concerned, the modern taxonomist readily agrees (Mayr, 1952) that species are not "discovered" but "invented." The principal problem of "validation" at this level is to establish the "existence" of a category, which means, essentially, that other investigators can distinguish the same grouping if they follow directions for finding it.

---

*Reprinted with permission from Bruner, J. S., Goodnow, J. J., and Austin, G. A., A Study of Thinking, (C) 1956 by John Wiley and Sons, Inc., New York, N.Y.

> It is at the next level of taxonomy, beta taxonomy, that the criterion of consistency becomes critical, for now the task is to order the bewildering array of species into a system of classification. Whether or not one's grouping of a series of species into genera and then into a phylum is "valid" or not depends upon whether the properties of the grouped species are consistent with one's conception of evolution. The correctness of the grouping "elasmobranchs" as distinguished from "teleostean fishes" depends upon whether the defining morphological properties of the two classes fit a consistent pattern of development as formulated in evolutionary theory. Teleostean fishes are more "evolved" or "higher than" sharks because of the consistent differences in skeletal system, renal system, circulatory system, etc. What makes the classification valid is that it is explicable in terms of a more general theory about the changes in morphology and physiology that characterize the evolution of animal life. If one would seek to establish the validity of considering a certain group of animals as constituting a phylum, the test would be by consistency with the requirements of the theory governing classification not only of this new phylum but of phyla generally," (4, p. 19).

As the authors cited above point out, test (validation) by consistency requires a context by which categorization can be judged. In the case of alpha taxonomy, the context is a procedural context, while in beta taxonomy, the context is theoretical or nomological. The question in the present study is whether test by consistency is applicable to deciding the extent to which a problem has been solved.

Test by consistency may be applied to problem solution on two levels roughly analogous to the alpha and beta levels of taxonomy. At the lower or alpha level one would expect the problem solution to fit into the context of the problem or, stated another way, the solution to the problem would be expected to be relevant to the informational cues provided by the problem context. If the task is for the teacher to diag-

nose the significance of the arithmetic errors made by a pupil on an arithmetic test, then the test provides a definite body of information within which the teacher works and to which his problem or task solutions will hold greater or lesser relevance. Presumably any teacher who examined such an arithmetic test could make the same diagnosis as any other teacher, if the procedures by which the diagnosis is to be made were outlined in sufficient detail.

It is apparent that the utility of a consistency criterion in the sense outlined (alpha level) depends upon the clarity of the informational cues. If the cues are very ambiguous, this criterion of consistency would be of little value.

At the beta level, one would expect problem solutions to be consistent not only with the information provided in the problem situation, but also with some type of empirical evidence, rationale, theory or nomological net which sets forth how the solutions to a given class of problems can be determined. For example, teachers who create original units of study in social studies frequently visit school and public libraries to obtain reading materials for children to use in the unit. The task of the teacher in such cases is to choose books and pamphlets which (i) are relevant to the unit and (ii) can be independently read by the children. Assuming that the teacher can find relevant material, the next problem is to determine whether it can be read by the children in his group. If he can find readable material he can proceed with the

unit. If he cannot, then he must either modify the means of getting information within the unit or else change the unit. Thus the possibility which first appears, namely taking back to the class what relevant books and pamphlets he can find and testing them out with the children to see whether they are readable, is not a fully satisfactory solution, even though an ultimate criterion for the readability of the material by his class is involved, since it is necessary to begin the unit before determining whether the relevant material available is readable.

The next possibility is to appeal to an external rationale for determining readability. Upon the supposition that the teacher teaches in grade 4 or 5, he might use the Dale-Chall formula (7) for predicting readability, selecting among his materials on the basis of sentence length and difficulty of vocabulary. With the information about readability provided by the Dale-Chall formula plus information about the reading ability of his class, the teacher can resolve the problem on the basis of consistency with an external rationale and the information available in the problem situation.

Suppose now that an experimenter confronts a teacher with a situational problem similar to the one just described, so that the teacher is actually choosing among materials to use with a fourth grade class whose reading comprehension levels are objectively established. In determining whether the teacher's choices are appropriate, the experimenter may also appeal to the external rationale provided by the Dale-

Chall formula, and thus use a consistency criterion determining problem solution. Moreover, he may examine the teacher's processes in solving the problem on the basis of the extent to which his search through the material is guided by such criteria of readability as sentence length and difficulty of vocabulary.

Whether a consistency criterion at the beta level is useful for determining when a problem is solved depends primarily on the relevance of existing rationales and theories to specific teaching problems. When a rationale or theory is untested, ambiguous or vague, as a great many "theories" in education are, it is useless as the basis of the application of a consistency criterion.

Procedure 3. Test by consensus.

The validation of a category by consensus is resorted to, according to Bruner, Goodnow, and Austin, when "...defining properties are vague and disjunctive..." or, "...when the criteria to be used are ambiguous in nature..." (4, p. 20). The way in which a consensual procedure may be used in deciding the degree to which a problem is solved is rather clear cut; a criterion group is selected and the problem solving performance of this group is then taken as the criterion for problem solution. For instance, in a task involving the teaching of borrowing in subtraction, one might obtain the performance of experts in teaching borrowing and evaluate problem solution of teachers on the basis of this

performance, or one might ask a group of elementary school teachers to perform such a task and then take the modal responses of this group as the criterion for problem solution. The use of consensual procedures for problem solution will lead to circularity in the projected series of studies, as will presently be shown.

Categorizing procedures and operational definitions of problem solution.

Which of the three criteria above (the criterion of effects, a consistency criterion or a consensual criterion) are chosen for categorizing the degree to which a problem is solved, ultimately determines what is meant by problem solving and directly bears on the operational meaning of proficiency in solving problems in teaching. Because each criterion for validating problem solution differs from the others, it is apparent that in a given instance "problem solution" may have more than one meaning, viz., as many meanings as there are criteria by which to validate the categories of problem solution. To utilize too many different criteria in determining problem solution, or to have too many operational meanings of "problem solving" would result in ambiguity. One purpose, then, of ordering criteria is to avoid confusion attributable to ambiguity. A second purpose is to avoid circularity in defining what is meant by "problem solving proficiency."

The avoidance of circularity is, perhaps, the more important matter in deciding upon a procedure by which to validate problem solution.

Observing the working purpose toward which the projected series of studies are directed, it is apparent that it would be circular to select a criterion group of teachers whose problem solutions would be considered as the "best" ones. If this procedure were used it would necessarily be the case that the group which one is seeking to isolate by means of the studies, viz., proficient solvers of teaching problems, has already been identified before the study begins. Moreover, the same objection may be raised in using experts as the criterion group since such a procedure would mean ipso facto that the experts are the best problem solvers. Whether the experts are experts because they are good solvers of teaching problems or for some other reason becomes the focal question in this case. To answer this question experts must presumably solve problems to which they have not themselves determined the answers. Meeting the latter condition implicitly entails eliminating a consensual criterion for problem solution. It is thus the case that a consensual criterion or the use of criterion groups would be of no use in the validation of problem solutions.

This does not mean, however, that experts cannot be used to determine whether the solution to a given problem can legitimately be inferred from a given rationale, which is simply making certain that a consistency criterion actually applies to a given problem. Moreover, exclusion of a consensual criterion as a means of determining

problem solution does not exclude the use of a consensual procedure in determining criterion groups for the empirical validation of problems, as will later be discussed.

The criterion of effects for the validation of problem solution successfully avoids circularity since there is no a priori assumption about the group which will produce these effects. However, the utility of this criterion is slight insofar as there is at present no means by which to control the stimuli which constitute actual classroom problems and hence no possibility of a uniform basis of comparisons between teachers or between occasions for the same teacher.

Of the three procedures for validating problem solution, then, only the consistency criterion remains. The application of this criterion at either the alpha or beta level depends upon the use of "artificial" or constructed problems in which the cues can be controlled. The utility of the alpha level criterion depends primarily on the degree to which evidence for one solution as opposed to its alternatives is presented within the problem. On the basis of the information or evidence presented, then, the subject can determine which among several alternative solutions is probably the best.

The use of the beta level of the criterion of consistency is free of circularity as long as either the research evidence or the research evidence plus particular rationale are sufficient to determine what the

solution to a given problem should be. However, because much of the "theory" in contemporary education consists of a "point of view" supported with modest evidence, there is an implicit danger of confusing subjects who disagree with a particular point of view with subjects who show little facility in solving teaching problems. The utility of the beta level of the criterion of consistency is obviously limited by the extent of theory development in the major areas of instruction in the elementary school.

The criterion to be employed for valid problem solution, then, is the consistency criterion. Some problems will be "correctly" solved when the solution is consistent with the information presented in the problem. Other problems will be correctly solved when the subject presents a solution consistent not only with the evidence in the problem, but also with some sort of rationale which transcends that particular problem.

Operationally defining proficiency.

The procedures by which problem solution is defined are the primary determinants of the operational definition of "proficiency" in problem solving. That is, since problem solution on the alpha level of a consistency criterion is defined by the evidence for a given solution within the problem, a proficient problem solver is necessarily one who discriminates and correctly manipulates the various items of information

to reach the solution. The same is true on the beta level with the additional criterion that the classifications and manipulations be consistent with an external rationale. There is, however, a further meaning of proficiency, namely, how well a person performs in comparison to other persons when the situation in which performance is recorded is held constant. From the viewpoint of operationism, "proficiency" in the latter sense derives its operational meaning from the sampling procedures used to determine the composition of the comparison group or groups.

The sampling procedures used to select the comparison groups in the projected studies are of importance primarily because these procedures have a direct bearing on which particular groups of elementary school teachers are identified as those who enjoy a high level of problem solving proficiency and who have continued to improve in proficiency. For instance, the best problem solvers in some school systems may be only as proficient as the worst problem solvers in other school systems. It is in turn quite feasable that the characteristics of the "most proficient problem solvers" could change according to the type of sample drawn. Because of the great diversity of teaching situations and criteria for teacher selection over the nation, the matter of sampling is a complex one, and can lead to conflicting results. It is therefore of the utmost importance that sampling procedures be an implicit part of the operational definition of proficiency.

Up to this point, proficiency has been discussed primarily on what may be called a "product" criterion, that is, the investigators have been centrally concerned with whether a subject can or cannot produce a solution as defined by a given criterion. An ultimate criterion, a consensual criterion and a consistency criterion, might have been used to define the solution, although all but the consistency criterion were rejected. It is apparent, however, that as well as talking about whether a subject solves a problem one can talk about how he goes about achieving a solution. In the latter case one is concerned with the problem solving process. How a problem is solved is, of course, bound up with whether it is solved, i.e. the product is the end of but still part of the process. Nevertheless, it is possible to talk about criteria on which the problem solving process is examined, or a process criterion, as well as talking about a product criterion. In the present rationale there are two types of criteria employable in examining behavior on a process criterion. One of these is the beta level of the criterion of consistency, as earlier suggested. Because much theory in education deals with the methods of teaching, the procedures a teacher uses in performing some tasks, i.e. those in which a teaching method is focal, may be judged on the basis of their consistency with accepted theory, whenever such theory is sufficiently exact.

The second type of criterion is empirical and is composed of the response variables which may be identified during the problem solving

process and subsequently quantified. The identification of such variables presupposes a situation in which the problem solving behavior of the teacher can be forced into the open and recorded. Individual teachers working on individual problems in the presence of an experimenter with apparatus for registering responses are necessary. Such conditions, however, are not necessary to assess problem solving proficiency on only a product criterion. When only a product criterion is used for assertions about problem solving proficiency, problems may be presented in a form similar to that of a group test. When both product and process criteria are used, problems must be presented under conditions similar to those of a problem solving experiment in a psychological laboratory.

Whether the experimenter chooses to use only a product criterion or both product and process criteria depends on whether he wants to obtain the responses of many persons in relation to a few dependent variables, as when using a product criterion, or whether he wants to obtain the responses of a few persons in relation to a wide array of dependent variables, as when both product and process criteria are used. Both of these approaches are used in the studies subsequently to be reported.

A number of descriptive dependent variables are of interest in examining the problem solving processes of teachers on teaching problems. Not all of the variables can be obtained from any one problem,

but rather may be obtained from an array of problems. Among these variables are the following: (i) time consumed by a subject in solving a problem; (ii) amount of information a subject requires beyond that which is necessary and sufficient to solve the problem; (iii) number of errors the subject makes--the blind alleys he enters; (iv) number of corrections of previous responses made by a subject; (v) position of corrections in the sequence of the subject's responses; (vi) number of relevant hypotheses which the subject can array and (vii) the strategy the subject uses in constructing categories. How many dependent variables might be dealt with beyond those mentioned depends on the cleverness of the experimenter in designing problems and apparatus to record responses. However, the focal point in identifying such dependent variables is to obtain as many indices as possible which bear on the amount and kind of non-solution achieving behavior emitted during the course of problem solution. The greater the array of non-solution achieving behavior emitted by a subject for equivalent degrees of problem solving, the less proficient the subject is as a problem solver.

In addition to using descriptive dependent variables as indices of non-proficient processes in problem solving, it is a matter of concern that the independent variables of which such dependent variables are a function be explored. The purpose in exploring such independent variables is primarily to be able to determine what characteristics of

teachers are associated with problem solving proficiency. A second matter of concern is to explore the ways in which significant independent variables can be manipulated in order to increase problem solving proficiency. From the viewpoint of teacher education a knowledge of those variables which can be manipulated to increase the problem solving proficiency of teachers would be of great utility. The achievement of such knowledge, however, is obviously in the future and is contingent upon the acquisition of somewhat more basic information.

Summary, Problem 2

Many of the problems elementary school teachers are called upon to solve do not have perfectly determinant solutions. It is therefore necessary to order criteria by which the several thinkable solutions may be judged. There are three main types of criteria which might serve this purpose, (i) an ultimate criterion, (ii) a consistency criterion, (iii) a consensual criterion. These criteria were discussed and it was concluded that only a consistency criterion is useful within the present purposes.

Problem solving proficiency may be defined upon two criteria, a product criterion, i.e. whether a problem is solved, or a process criterion, i.e. how a problem is solved. Proficiency on a product criterion may be defined by a consistency criterion on either or both of two levels--the alpha level, which requires that problem solution

be consistent with the information or evidence provided within the problem, or the beta level, which requires in addition that problem solution be consistent with a rationale external to the problem. Proficiency on a process criterion may also be defined by a consistency criterion at both the alpha and beta levels since a process criterion encompasses a product criterion. However, a process criterion for problem solving proficiency may be defined as well by those quantifiable, dependent variables which serve as indices of the amount and kind of non-solution achieving behavior emitted by the subject. A general summary of the criteria and their relationships may be observed in Figure 2.

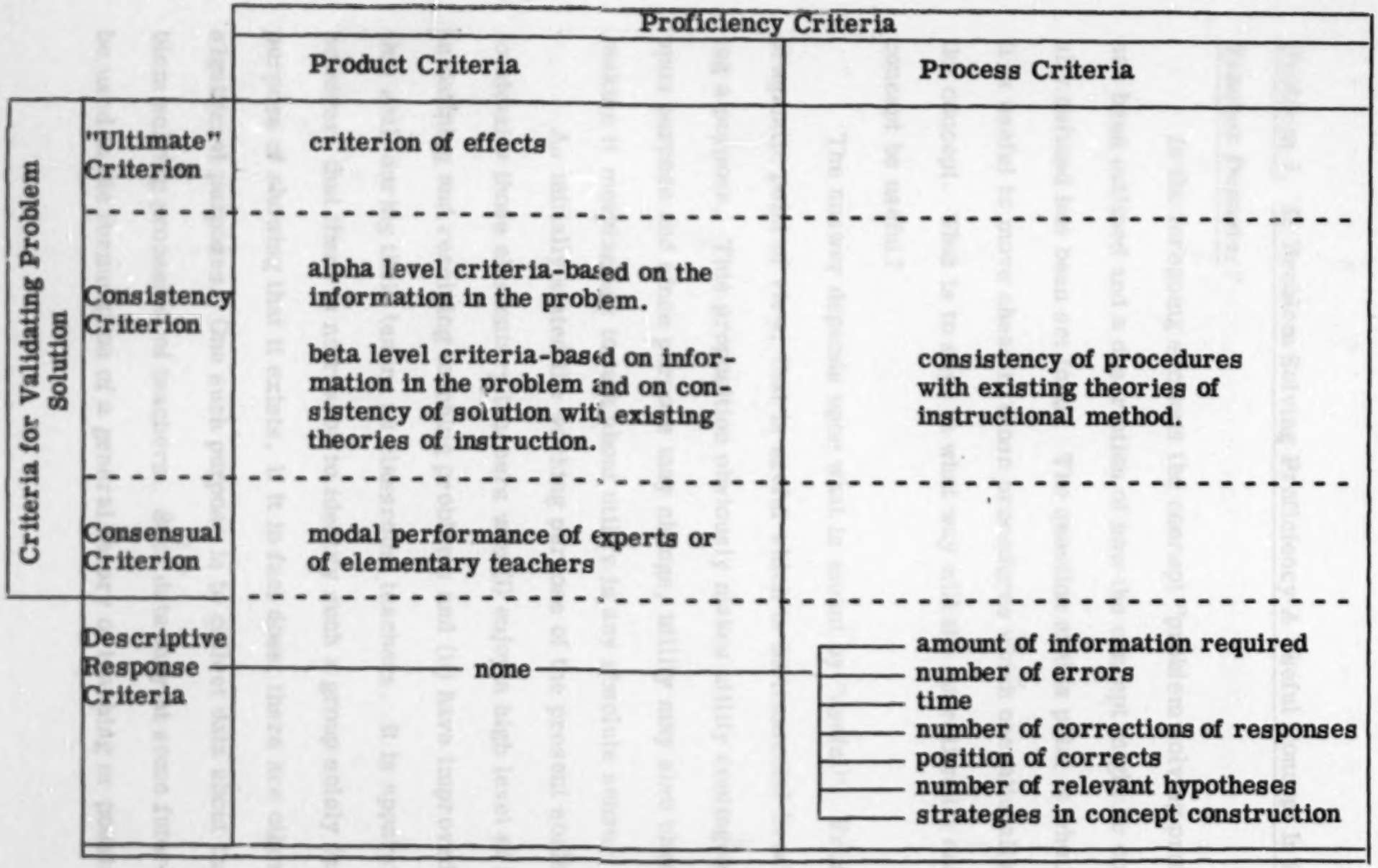
Proficiency Criteria
Product Criteria
Process Criteria
Criteria for Validating Problem Solution
"Ultimate" Criterion
criterion of effects
Consistency Criterion
alpha level criteria-based on the information in the problem.
beta level criteria-based on information in the problem and on consistency of solution with existing theories of instruction.
consistency of procedures with existing theories of instructional method.
Consensual Criterion
modal performance of experts or of elementary teachers
Descriptive Response Criteria
none
amount of information required
number of errors
time
number of corrections of responses
position of corrects
number of relevant hypotheses
strategies in concept construction

FIGURE 2 A Summary of Criteria for Problem Solution

Problem 3. Is Problem Solving Proficiency A Useful Concept In Studying Teacher Behavior?

In the foregoing sections the concept "problem solving proficiency" has been outlined and a description of how the concept might be operationally defined has been set forth. The question at this point is whether it is useful to move ahead to those procedures which operationally define the concept. That is to say, in what way will the operationally defined concept be useful?

The answer depends upon what is meant by "useful". From a pragmatic point of view, that is useful which is instrumental in achieving a purpose. This proposition obviously makes utility contingent upon purpose and since purpose may change, utility may also change, making it meaningless to talk about utility in any absolute sense.

As initially stated the working purpose of the present studies is to identify those elementary teachers who (i) enjoy a high level of skill in defining and resolving teaching problems and (ii) have improved in this skill during their tenure as classroom teachers. It is apparent, however, that there is no reason to identify such a group solely for the purpose of showing that it exists, if it in fact does; there are other more significant purposes. One such purpose is to collect data about the problem solving processes of teachers. Such data may at some future time be used in the formulation of a general theory of teaching or possibly

incorporated into a theory of problem solving. This purpose is essentially that of doing alpha level taxonomic research in teacher behavior and in problem solving behavior. Within this purpose are a number of goals, including (i) isolating the response variables which distinguish proficient from non-proficient problem solvers, (ii) isolating the independent variables of which the response variables are a function and (iii) discovering how independent variables can best be manipulated to maximize problem solving proficiency.

A second of the more significant purposes arises out of the discrepancy between the working purpose and the operations adopted for achieving this purpose. In order to obtain the proper control over the conditions under which problem solving behavior is to be examined, it has been assumed to be necessary (i) to control the problems used so that they are functionally identical for all subjects, and (ii) to control the stimulus context in which the problems appear so that there is high probability that the subject is responding to the stimuli constituting the problem the experimenter is interested in, rather than some other stimuli or some other problem. In gaining this kind of control it is necessary to use problems which are not identical to the problems teachers face in the classroom.

If one is going to identify a group of teachers who are proficient solvers of problems encountered in teaching, it is obviously necessary

to show that some sort of relationship holds between solving problems under laboratory or test conditions and solving problems under natural or classroom teaching conditions. The question here is one of demonstrating the relevance of behavior obtained under one set of conditions to behavior obtained under another set of conditions, and is essentially a question of demonstrating the validity of a measure or set of measures.

The validation of laboratory problems.

In a superficial appraisal of the problem, the most direct way of showing that the laboratory behavior of teachers is relevant to their classroom behavior is to perform the laboratory measurement, then send the teachers back to the classroom and measure them there. The predictor-criterion relationship then seems adequate and it is, except for one previously mentioned problem. It is difficult to control the circumstances which surround the teachers' classroom behavior, i.e. the conditions in the criterion vary. To predict an uncontrolled criterion from a controlled predictor is obviously unsound. The problem of validation must be approached by less direct methods.

Of the currently available methods of validation, construct validity, as elaborated by Cronbach and Meehl (6) and criticized by Bechtoldt (2) and intrinsic validity, as elaborated by Gulliksen (9) bear relevance to the problem at hand. Of these two methods of validation, intrinsic validity appears to be the more parsimonius and most relevant to the validation problem with which the authors are concerned.

The central concern in establishing the intrinsic validity of a measuring device is to develop criteria such that the class of behaviors measured by the device are very clearly specified by the nature of, and relationships between, the criteria which the device will predict. The focus in validation is fully on the criteria to be predicted. The central question to be dealt with in this section of this paper is, "What kinds of criteria can be developed by which to validate constructed teaching problems?"

In his discussion of intrinsic correlational validity, Gulliksen suggests as one criterion the relationship between training and test performance (9, p. 513). In intrinsic correlational validity this criterion consists of demonstrating that training on the items of a test or on constructed problems results in improvement in criterion performance. This procedure is not highly workable in the present case since the criterion measure, i.e. classroom problem solving performance, cannot be controlled. However, the reverse process is workable, i.e. demonstrating that classroom experience related to test problems will result in an increment in test performance. Whether this reversal of predictor-criterion relationships is permissible will be briefly examined.

Demonstrating that training on test items improves criterion performance is one way of asserting that positive transfer occurs from the training trials on a particular task or set of tasks to test trials on a

somewhat different set of tasks. In the reverse case the training trials and test trials are simply interchanged so that what was previously a predictor is now a criterion and what was previously a criterion is now a predictor. Again positive transfer would be expected to occur. If it did not, an inconsistency would be involved since an increase in "criterion" performance (now the predictor) would not be reflected in test performance (now the criterion), while an increase in test performance (now the predictor) would result in an increase in criterion performance, (now in fact the criterion). It is difficult to imagine a set of circumstances in which such inconsistent relationships would actually hold. This is to say, then, that when a test is intrinsically valid, positive transfer will occur regardless of whether the training trials deal with the test items or with the particular experiences which constitute the criterion.

Experimental designs for validation on a transfer criterion.

In showing that constructed problems in teaching are valid on the criterion of positive transfer from training to test trials, at least two experimental designs are operable. In the most simple case, the first is a longitudinal design with two groups, both drawn from the college graduate population. Group A, the control, would be drawn from the non-education curricula of each college involved; Group B, from the education curriculum. Both groups would be tested on the constructed

teaching problems near graduation and then tested yearly thereafter. The predicted results, if the test is intrinsically valid, are that those persons who enter teaching and instruct in those areas to which the problems are relevant, i.e. arithmetic, reading, etc. in the elementary school, will, up to some asymptote, increase in performance, while the control group will remain essentially constant in performance.

Within this design there is a lack of control of the quality of the experiences of the teachers group. The kind of experience, i.e. type of training trials, the teacher undergoes is obviously a source of variation. However, the fact that there may be various degrees of transfer within the teacher group is less important from the viewpoint of the validity of the instrument than whether there are changes in the scores of the control group, which presumably should be having few or no experiences directly related to teaching problems. If the test is valid on the criterion being discussed, the means and variances of the control group should remain essentially the same or perhaps decrease slightly, while the measures of central tendency of the experimental group should increase with a distinct possibility that the variances will also change significantly.

The second design, in the most simple case, is based on the cross-sectional observation of two groups, teachers and non-teachers. The control group is composed of persons who do not teach, have never

taught nor have ever been instructed in teaching arithmetic, reading, etc., in the elementary schools. The experimental group, on the other hand, has been trained to teach and, in fact, instructs in those areas in which the constructed tasks fall. If the constructed tasks are valid, one would expect the mean performance of teachers to exceed that of non-teachers.

While this design is much easier to carry out than the first, it has a definite limitation upon the supposition that the experimental group outperforms the control group, for there is no definite information available within the design to indicate that this result is attributable to training rather than to selective factors. This is to say that this design will permit rejection of the null hypothesis but will not indicate which of two competing alternative hypotheses to accept.

The necessary and sufficient conditions for meeting the intrinsically valid criterion in question is to reject the null hypothesis and accept the alternative hypothesis that the differences are a result of training. The first design meets the necessary and sufficient conditions, (i.e. will show whether there are differences between groups and show whether these differences are attributable to training), the second design meets only a necessary condition (i.e. shows only that there are differences between groups) and is thus the weaker. The latter remains an operable design, but the validity of the constructed problems might be questioned until both the necessary and sufficient conditions of the

positive transfer aspect of the intrinsically valid criterion are met. The place of the weaker design is in pilot study work since if the necessary condition cannot be met there is not much point in worrying about meeting both the necessary and sufficient conditions.

Both the designs presented are given in most simple form for purposes of illustration, but it is apparent that both might be expanded into factorial designs from which the information relevant to validation may be extracted, as well as information concerning other sources of variation, and interactions of interest to the experimenter.

Establishing additional criteria for intrinsic validity.

The intrinsically valid criterion described in the two designs above must be met by a set of constructed problems to demonstrate their relevance to teaching. This criterion has, however, at least two other aspects which do not focally appear in the above designs. One of these deals with criterion groups of teachers and the other with learning ability.

If it is supposed that one of the two designs mentioned is carried through and the constructed problems are valid on the transfer criterion utilized, a question will remain concerning how certain groups of teachers perform.

One of these groups consists of teachers who might be shown, by the measurement of pupil gain, to produce unusually good results

year after year regardless of the conditions under which they teach. While pupil gain, measured at the end of the teacher instrumental responding, is not an ultimate criterion for problem solving proficiency, and while it is not possible to hold conditions constant between teachers, teachers who can consistently produce unusual gains as conditions vary, clearly form a relevant criterion group upon which the validity of constructed teaching problems can be tested. Failure to differentiate this group from a random sample of teachers by means of constructed teaching problems would clearly provide damaging evidence against the intrinsic validity of the problems.

A second criterion group of teachers consists of those who have been formally employed as critic, demonstration or consulting teachers because of recognized skill in teaching. The relevance of this group does not rest upon having demonstrated unusual problem solving skill upon a specific criterion, but rather on the assumption that the poorer solvers of teaching problems will tend to have been eliminated from the group in the general sorting-out process which precedes employment as a critic, demonstration or consulting teachers. A random sample from the population of such teachers would be expected to yield higher mean performance than a random sample from the general population of teachers because the lower end of the distribution of problem solving skill will be truncated in the select group.

If, as in the longitudinal, two-group design set forth, problem solving proficiency is expected to improve, up to some asymptote, with teaching experience, it is the case that learning from teaching experience is implicitly involved in the transfer criterion and that learning ability is therefore a significant source of variation in the improvement of problem solving proficiency. Other things being equal, then, differences in learning ability should produce differential effects in problem solving performance. The latter proposition implies that performance will also be functionally related to intelligence. It is clear, however, that the extent of the relationship between problem solving proficiency and intelligence will be determined by the extent to which the particular intelligence test used is in fact a measure of learning ability. As Gulliksen points out, (9, p. 515) the relationship between intelligence test scores and learning ability under laboratory conditions is characteristically of modest magnitude. Nonetheless, consistently negative or zero correlations between intelligence test scores and problem solving proficiency scores would weaken confidence in the validity of the instrument used to measure problem solving proficiency in teaching.

Still another type of criterion of interest in establishing the relevance of performance on constructed problems to performance on problems arising in the classroom lies in the relationships between the scores on constructed problems and scores on tests other than intelligence. The problem may be seen as one of accounting for too much

of the variance in constructed problem scores by means of tests which bear no apparent relevance to the solving of classroom problems. For instance if 50% of the variance of the scores on constructed teaching problems is accounted for by a mechanical aptitude test which reliably predicts success in repairing automobile engines, confidence in the validity of the former instrument is seriously weakened. There is no apparent rationale by which such a close relationship between solving problems in teaching and mechanical aptitude can be explained. The identification of such discrepant relationships obviously must be done on a strictly empirical, exploratory basis, since it is precisely those variables that seem logically extraneous which, if shown to be empirically relevant to test performance, serve as indicators that the validity of a test is spurious.

Establishing the concept of a professional teacher.

A third purpose for which the concept of problem solving proficiency in teaching is useful is as a means of arriving at a criterion for identifying a "good teacher" or in operationally definable terms a professional teacher. Whether the concept is useful for this purpose depends upon the acceptance of three subordinate criteria, one of which is analytic, one of which is value-laden, and one of which is "empirical."

The analytic criterion is the proposition with which this paper began, namely that the term "good elementary school teacher" strictly entails the term "good solver of problems within the domain encompassed

by the objectives of the elementary school." The latter term is not the only term entailed by "good elementary teacher," but it is one of them. The intent of this proposition is, of course, that it is logically inconceivable that those who are defined as elementary school teachers cannot also be defined as those who solve problems dealing with the education of the young. Clearly the proposition "This young lady is an elementary school teacher and she solves no problem involving the education of young children" involves a contradiction and is absurd. However the proposition "This young lady is an elementary school teacher and she guides the education of the young" is redundant, but not absurd and points toward the fact that out of all the terms that might describe what it is a teacher does, the authors have chosen only one, problem solving.

There are a number of reasons for this choice. The first, as previously pointed out, is heuristic and is an attempt to break away from the concept of "teacher effectiveness" which has had, to put the matter mildly, a blinding effect on research in this area. A second reason is that by viewing teaching as problem solving activity it is possible to identify it as related to instrumental activity and particularly to conceptualize teaching as involving much instrumental learning. In turn in asking such a question as "why do some teachers perform better than others?", it is possible to search for answers among the variables which

effect instrumental learning rather than dust off a favorite arm-chair theory. Finally, by identifying teaching with problem solving activity it becomes fully conceivable that a great deal of the behavior relevant to teaching can be examined under laboratory conditions where some control over variables can be exercised.

The value-laden criterion consists of the objectives by which the domain within which elementary school teachers should work is defined or delimited. Elementary school teachers obviously must work within certain limits; they are not charged with healing the sick, ministering the benefits of religion, or dispensing legal advice. There is, however, a difference between asserting that teachers do work within a particular domain and asserting that it is desirable that teachers work within a particular domain. The assertion that teachers do work within a particular domain or do in fact work for certain objectives is an assertion open to empirical verification, whereas the assertion that it is desirable that teachers work toward particular objectives is not open to verification by standard empirical procedures. There has been a great deal of difficulty over this point in research in teacher effectiveness with a strong tendency to act as if the value judgments made about teaching are not centrally significant sources of variation in determining who is effective. Yet, it is the very discrepancy between what is and what a rater or group of raters, a test constructor or group of test construc-

tors, including constructors of tests of pupil achievement, think ought to be, which is, in fact, the actual criterion for statements about teacher "effectiveness."

The difference between what is verifiable and what is valued obviously cannot simply be brushed aside and ignored since this procedure permits value judgments to creep into criteria in devious ways. In making a frontal attack on the aspects of teaching involving value judgments, the authors utilize two assumptions. The first is that the domain of objectives for which teachers ought to work overlaps the domain of objectives for which they do in fact work. If this assumption is not made, there are no grounds for supposing that positive transfer will occur from teaching experience, which is composed of what teachers do, to constructed teaching problems, which are based on what it is thought teachers ought to be able to do. It is clear, moreover, that the difference between what teachers do and what they ought to do, as defined by the discrepancy in two sets of objectives, must be a source of variation in their performance on constructed problems in teaching. For instance, if the objectives of a teacher of arithmetic in the elementary school deal only with manipulating numbers, he may become very proficient in solving problems in teaching the manipulation of numbers but if he is asked to solve problems dealing with understanding mathematical meanings of arithmetic he may be much less proficient. The

source of variation in this instance is not the teacher's learning ability, but rather the kinds of objectives which he has learned to respond toward and which are presumably the objectives he holds to be of value.

Now, while it may be asserted that the discrepancy between what a teacher does and what he ought to do is a source of variation in the performance of teachers, the assertion remains empty unless the source of variation can be quantified independent of the teacher's problem solving performance. There are three ways in which this might be done. The first is observation of the kinds of goals a teacher actually works toward. This procedure has apparent limitations because of the time involved in taking adequate individual observations and because of difficulties in taxonomizing and quantifying observations. A second procedure is to examine the stated objectives of the school system in which the teacher works. This procedure is weak in that it assumes that school systems have bothered to state their objectives and that if they have, the teachers actually work toward the stated objectives. The third procedure is to construct a measuring device by which to assess the kinds of behavioral outcomes in pupils a teacher believes to be valuable. While the construction of such a device would involve complex validation procedures, its utility in the projected series of studies and for other studies can hardly be doubted. A currently available instrument which, while peripheral, nevertheless takes the general direction

of the instrument the authors have in mind is the Minnesota Teacher Attitude Inventory (5).

The second assumption in making a frontal attack on the aspects of teaching involving value-judgments is that the objectives which have been published by various leaders in the profession within recent years may be taken as those objectives which it is desirable that teachers be able to bring about. It is not likely that such objectives are all equally acceptable to all the members of the teaching profession, but this disadvantage is less focal than the fact that objectives determined by reputable professional persons such as Brownell (3),. Greene and Gray (8), Havighurst (10), and Kearney (12) can hardly be viewed as arbitrary objectives and that by utilizing such objectives it is possible to provide an anchoring point from which systematic research can begin.

Both the analytic and the value-laden criteria set forth are essentially definitional criteria. These criteria may be regarded as "professional criteria" insofar as the ability to solve problems generated by certain objectives arising within the profession acts as a standard by which persons can be categorized or ordered within the profession. As soon as objectives are found and problems constructed, the standard may be operationally defined by the distribution or distributions which result from performance on them. When these procedures are finished, a concept has been defined, i.e. problem solving proficiency. This concept has operational meaning but it may not have significance (2, p. 623).

In order to have significance or utility, the operationally defined concept must hold some sort of empirical relationships. In the present case these relationships are those described under the discussion of intrinsic validity which forms the third and the empirical criterion by which the concept of a professional teacher is defined.

At this point it is important to distinguish very clearly between problem solving proficiency among teachers, which is the concept that has operational meaning, but whose significance depends on the kind of empirical relationships it holds (and it might hold a great many which the authors have not thought about) and the concept of a professional teacher which involves, besides the concept of problem solving proficiency, the demonstration of those empirical relationships necessary to intrinsic validity. When these relationships can be shown to hold, then the operational definition of a proficient solver of constructed teaching problems, as defined by the procedures used, will be the same thing as a professional teacher. When these relationships do not hold, a proficient solver of constructed teaching problems will be a proficient solver of constructed teaching problems--and no more.

Finally it is only fair to recognize that the notion of a "professional teacher" involves a great deal of meaning which the authors simply cut out in the operational definition of that term. A good part of this meaning is ably discussed by Lieberman (13). The fact remains, however, that in order to be useful in the growth of educational technology, the extended meaning of "professional teacher" must be translated from words into operations.

SUMMARY

The function of the present report is to set forth a systematic rationale by which subsequent empirical studies of skill in elementary school teaching may be guided. The working purpose of these studies is to identify a group of elementary teachers who show high proficiency in solving problems in teaching upon a given occasion, and who increase to an asymptote in proficiency as teaching experience increases. The identification of such a group of teachers, however, is not an end in itself. There are two ulterior purposes. The first is to acquire data about the problem solving processes of teachers and about the independent variables which effect these processes. The second is to show that the operations which lead to the identification of this group will permit the emergence of an operationally defined concept, the professional teacher. The operational meaning of this concept rests upon problem solving performance, but its significance rests upon the relationships between problem solving performance and a matrix of relationships which constitute the intrinsically valid criteria on which to make assertions about teaching skill. One of these criteria entails that positive transfer occur from teaching experience to problem solving performance. This criterion further entails that learning ability of the teacher and the educational objectives the teacher selects to respond toward will be significant sources of variation in the magnitude of transfer from

teaching experience to problem solving performance. A second of the criteria entails that selected groups within the profession, i.e. (i) teachers who consistently produce unusual pupil gains, and (ii) critic, demonstration or consulting teachers will outperform random samples of teachers. These criteria do not exclude the use of other operationally definable criteria which are consistent with those criteria set forth, as they become apparent to the researcher.

In order to achieve an operational definition of problem solving proficiency and demonstrate that the concept of a professional teacher is significant, certain assumptions have been made and certain conditions of attaining the concept of a professional teacher stated. It was first assumed that teachers are professional persons whose function is to define and resolve the problems which arise because certain educational objectives are projected for pupils to attain. Second, it was assumed that as professional persons, it is desirable that teachers assume and be capable of attaining those objectives asserted as desirable by the leaders of the profession.

Once these two assumptions, which indicate the grounds on which teaching problems arise, were made, it was then necessary to set forth the conditions under which it would be considered permissible to assert (i) that a teacher has solved a problem and (ii) that a teacher is proficient in solving problems. This was essentially the task of setting forth the types of criteria by which it may be determined whether a problem is solved, and then setting forth the criteria by which proficiency is defined.

Three types of criteria were set forth by which it might be determined whether a problem is solved. The first was an ultimate criterion. For an ultimate criterion to be applied to problem solution, the problem must be such that its solution can be unmistakably recognized as the solution. No ambiguity is involved. The effects of the teacher as measurable in pupil behavior have been held to be the ultimate criterion for whether a problem in teaching is solved. However, in order to use effects as an ultimate criterion it is first necessary to demonstrate that a particular pupil behavior is in fact attributable to a given teacher and it is necessary second to define an "effect" so that it is not ambiguous. Up to the present the criterion of teacher effects has not been able to meet those conditions. Whether they ever can be met is a focal question.

A second criterion set forth for deciding whether a problem is solved was a consistency criterion. A consistency criterion may be applied to problem solution on two levels. At the alpha level the question is whether the solution to the problem is consistent with the information available for determining the solution to the problem. Where there is little information or ambiguous information this criterion is not useful since there is basically nothing to be consistent with. At the beta level the question is whether the solution to the problem can be inferred from some type of empirical evidence, rationale, theory or nomological net. If such inferences are possible, then the criterion of consistency can be applied to problem solution.

A third type of criterion examined in relation to problem solution was the consensual criterion. This criterion is useful when the solution to the problem is so ambiguous that the only way to determine what the solution is is to get a community of opinion on the matter. The direct use of a consensual criterion for the determination of problem solutions was identified to involve circularity in the projected studies, although its use to confirm the legitimacy of an inference within a consistency criterion was not excluded. Of the three possible criteria to apply to problem solution, the consistency criterion was selected as most suitable for use in the projected studies.

Once the criteria upon which a problem may be considered to be solved had been established, the next concern was to establish the conditions and criteria for asserting that some persons are more proficient problem solvers than others. In order to make this assertion without equivocation, it is necessary to meet the condition that each problem solver is solving identical problems. To satisfy this condition laboratory problems rather than "natural" classroom problems must be utilized, otherwise the problems cannot be held constant. A second condition to be met is that when the comparative proficiency of a given teacher is asserted, the relevance of the comparison group to that teacher must be demonstrated.

Finally, proficiency must be defined not only by the degree to which a person solves standard groups of problems in comparison with

a relevant population, but also upon what the person does in attempting to achieve a problem solution. What the person does depends largely on the problem with which he is engaged. However, over a range of problems, the following response variables might be found to be relevant: (1) time to problem solution, (2) amount of information required beyond that which is necessary and sufficient, (3) number of errors, (4) number of corrections of responses, (5) sequence of responses and corrections, (6) number of hypotheses arrayed and (7) strategy in constructing categories. The extent of this list is largely contingent upon how astute the experimenter is in designing problems and apparatus by which to force the cognitive processes of teachers into the open where they can be recorded and quantified. In using these response variables as process criteria, the focal question is the extent to which any one of them or any combination of them will accurately represent the amount of non-solution-achieving behavior which is emitted by the subject while engaged in the problem. The greater the amount of non-solution-achieving behavior for equivalent degrees of problem solution, the less proficient the problem solver on a process criterion.

Having thus set forth the assumptions, criteria and conditions involved in operationally defining problem solving proficiency among teachers, it becomes possible to undertake the empirical studies which will determine whether or not the concept of problem solving proficiency is a significant concept.

Such studies have been undertaken during the past two years. Monographs that are in varying stages of preparation include the following:

A. Problem Solving Proficiency Among Elementary School Teachers. II. Teachers of Arithmetic, grades 3-6.

B. Problem Solving Proficiency Among Elementary School Teachers. III. Teachers of Reading, grades 2-5.

C. Problem Solving Proficiency Among Elementary School Teachers. IV. Teachers of Arithmetic, grades 3-6, a replication.

D. Problem Solving Proficiency Among Elementary School Teachers. V. An Analysis of Processes and Strategies.

E. Problem Solving Proficiency Among Elementary School Teachers. VI. Hypothesis Formation with Reference to Pupil Behavior.

REFERENCES

1. Barr, A. S. and Jones, R. E. "The Measurement and Prediction of Teacher Efficiency", Review of Educational Research, 28: 256-264, June, 1958.
2. Bechtoldt, Harold P., "Construct Validity: A Critique," The American Psychologist, 14:619-630, October, 1959.
3. Brownell, W. A., "The Evaluation of Learning in Arithmetic", Arithmetic in General Education, Sixteenth Yearbook of the National Council of Teachers of Mathematics, Bureau of Publications, Teachers College, Columbia University, New York, 1941.
4. Bruner, J. S., Goodnow, J. J. and Austin, G. A., A Study of Thinking, Wiley and Sons, New York, 1956.
5. Cook, W. W., Leeds, C. H. and Callis, R., Minnesota Teacher Attitude Inventory, Psychological Corporation, 1951.
6. Cronbach, Lee J. and Meehl, Paul E., "Construct Validity in Psychological Tests", Psychological Bulletin, 52:281-302, July, 1955.
7. Dale, Edgar and Chall, Jeanne S., A Formula for Predicting Readability, Bureau of Educational Research, Ohio State University, no date.

8. Greene, Harry A. and Gray, William S., "The Measurement of Understanding in the Language Arts", The Measurement of Understanding Forty-fifth Yearbook, National Society for the Study of Education, 1946, pp. 175-212.

9. Gulliksen, Harold, "Intrinsic Validity", The American Psychologist, 5:511-517, October, 1950.

10. Havighurst, Robert J., Human Development and Education, Longman, Green and Co., 1953.

11. John, E. R., Contributions to the Study of the Problem-Solving Process, Psychological Monographs, Vol. 71, No. 18, 1957.

12. Kearney, Nolan C., Elementary School Objectives, Russell Sage Foundation, New York, 1953.

13. Lieberman, Myron, Education as a Profession, Prentice-Hall, Englewood Cliffs, N.J., 1956.

14. Remmers, and all, "Report of the Committee on the Criteria of Teacher Effectiveness", Supplement, Review of Educational Research, 22:238-263, June, 1952.